Making healthcare smarter

IBM Content & Predictive Analytics
A Preparation for Watson: Seton Use Case

February 2012
IBM Software Solutions Group
Watson in Healthcare
Findings

Symptoms
- difficulty swallowing
- fever
- dry mouth
- thirst
- anorexia
- frequent urination
- dizziness
- no abdominal pain
- no back pain
- no cough
- no diarrhea

Family History
- Oral cancer
- Bladder cancer
- Hemochromatosis
- Purpura
- Graves’ Disease
- (Thyroid Autoimmune)

Patient History
- cutaneous lupus
- osteoporosis
- hyperlipidemia
- frequent UTI
- hypothyroidism

Medications
- Alendronate
- pravastatin
- levothyroxine
- hydroxychloroquine
- urine dipstick:
- leukocyte esterase
- supine 120/80 mm HG
- heart rate: 88 bpm
- urine culture: E. Coli

Diagnosis Models
- Renal Failure
- UTI
- Diabetes
- Influenza
- Hypokalemia
- Esophagitis

Confidence

• www.ibm.com/jstart
• solutions start here.
What if it is not known?... Not Watson

Question What is Known
Complement ICPA with Watson for Healthcare to get real time, confidence based answers with evidence based learning

Analyze and Visualize the Past
Understand trends, patterns, deviations, anomalies, context and more in large corpuses of historical clinical and operational information to reveal new insights

See the Present
Analyze and extract text from in-process documents or other information to find structured data errors... feed the results to other cases and systems

Predict the Future
Use predictive models and scoring to make more informed decisions through predictive and future scenario modeling
Seton Healthcare Family

Reducing CHF readmission to improve care

“IBM Content and Predictive Analytics for Healthcare uses the same type of natural language processing as IBM Watson, enabling us to leverage information in new ways not possible before. We can access an integrated view of relevant clinical and operational information to drive more informed decision making and optimize patient and operational outcomes.”

Charles J. Barnett, FACHE, President/Chief Executive Officer, Seton Healthcare Family

Business Challenge
Seton Healthcare strives to reduce the occurrence of high cost Congestive Heart Failure (CHF) readmissions by proactively identifying patients likely to be readmitted on an emergent basis.

What’s Smart?
IBM Content and Predictive Analytics for Healthcare solution will help to better target and understand high-risk CHF patients for care management programs by:

• Utilizing natural language processing to extract key elements from unstructured History and Physical, Discharge Summaries, Echocardiogram Reports, and Consult Notes
• Leveraging predictive models that have demonstrated high positive predictive value against extracted elements of structured and unstructured data
• Providing an interface through which providers can intuitively navigate, interpret and take action

Smarter Business Outcomes
• Seton will be able to proactively target care management and reduce re-admission of CHF patients.
• Teaming unstructured content with predictive analytics, Seton will be able to identify patients likely for re-admission and introduce early interventions to reduce cost, mortality rates, and improved patient quality of life.

IBM solution
• IBM Content and Predictive Analytics for Healthcare
• IBM Cognos Business Intelligence
• IBM solution services

© 2011 IBM Corporation
IBM Content and Predictive Analytics for Healthcare

The Seton CHF Readmission POC Solution

IBM Watson for Healthcare

Confirm hypotheses or seek alternative ideas with confidence based responses from learned knowledge*

Utilizing natural language processing to extract key elements from unstructured History and Physical, and Discharge Summary

Leveraging predictive models that have demonstrated high positive predictive value against extracted elements of structured and unstructured data

Content Analytics
- Natural Language Processing
- Medical Fact and Relationship Extraction (Annotation)
- Trend, Pattern, Anomaly, Deviation Analysis

Predictive Analytics
- Predictive Scoring and Probability Analysis

Dynamic Multimode Interaction

Providing an interface through which providers can intuitively navigate, interpret and take action

Monitor, Dashboard and Report (Cognos BI)

Question and Answer*

Custom Solutions

Health Integration Framework

Data Warehouse and Model
Master Data Management
Advanced Case Management
Business Analytics

Partners (HLI) Specialized Research

Raw Information

Unstructured Data
(EMR Clinical Documentation: History and Physical, Discharge Summary, Echocardiogram.)

Structured Data
(Cons Data, Admission History, Procedure History, EMR Clinical Events)
IBM Content and Predictive Analytics for Healthcare

The Seton CHF Readmission POC Solution

IBM Watson for Healthcare
Confirm hypotheses or seek alternative ideas with confidence based responses from learned knowledge*

Raw Information

Unstructured Data
(EMR Clinical Documentation: History and Physical, Discharge Summary, Echocardiogram.)

Structured Data
(Cost Data, Admission History, Procedure History, EMR Clinical Events)

IBM Content and Predictive Analytics

Content Analytics
- Natural Language Processing
- Medical Fact and Relationship Extraction (Annotation)
- Trend, Pattern, Anomaly, Deviation Analysis

Predictive Analytics
- Predictive Scoring and Probability Analysis

Analyzed and Visualized Information

Dynamic Multimode Interaction

Search and Visually Explore (Mine)
Monitor, Dashboard and Report (Cognos BI)
Question and Answer*
Custom Solutions

Health Integration Framework

Data Warehouse and Model
Master Data Management
Advanced Case Management
Business Analytics

Partners (HLI)
Specialized Research

© 2011 IBM Corporation
Seton Healthcare Family

Reducing CHF readmission to improve care

“IBM Content and Predictive Analytics for Healthcare uses the same type of natural language processing as IBM Watson, enabling us to leverage information in new ways not possible before. We can access an integrated view of relevant clinical and operational information to drive more informed decision making and optimize patient and operational outcomes.”

Charles J. Barnett, FACHE, President/Chief Executive Officer, Seton Healthcare Family

Business Challenge

Seton Healthcare strives to reduce the occurrence of high cost Congestive Heart Failure (CHF) readmissions by proactively identifying patients likely to be readmitted on an emergent basis.

What’s Smart?

IBM Content and Predictive Analytics for Healthcare solution will help to better target and understand high-risk CHF patients for care management programs by:

• Utilizing natural language processing to extract key elements from unstructured History and Physical, Discharge Summaries, Echocardiogram Reports, and Consult Notes
• Leveraging predictive models that have demonstrated high positive predictive value against extracted elements of structured and unstructured data
• Providing an interface through which providers can intuitively navigate, interpret and take action

Smarter Business Outcomes

• Seton will be able to proactively target care management and reduce re-admission of CHF patients.
• Teaming unstructured content with predictive analytics, Seton will be able to identify patients likely for re-admission and introduce early interventions to reduce cost, mortality rates, and improved patient quality of life.

IBM solution

• IBM Content and Predictive Analytics for Healthcare
• IBM Cognos Business Intelligence
• IBM BAO solution services
What Really Causes Readmissions at Seton

Key Findings

The Data We Thought Would Be Useful ... Wasn’t

- 113 candidate predictors from structured and unstructured data sources
- Structured data was less reliable than unstructured data – increased the reliance on unstructured data

New Unexpected Indicators Emerged ... Highly Predictive Model

- 18 accurate indicators or predictors (see next slide)

<table>
<thead>
<tr>
<th>Predictor Analysis</th>
<th>% Encounters Structured Data</th>
<th>% Encounters Unstructured Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ejection Fraction (LVEF)</td>
<td>2%</td>
<td>74%</td>
</tr>
<tr>
<td>Smoking Indicator</td>
<td>35% (65% Accurate)</td>
<td>81% (95% Accurate)</td>
</tr>
<tr>
<td>Living Arrangements</td>
<td>&lt;1%</td>
<td>73% (100% Accurate)</td>
</tr>
<tr>
<td>Drug and Alcohol Abuse</td>
<td>16%</td>
<td>81%</td>
</tr>
<tr>
<td>Assisted Living</td>
<td>0%</td>
<td>13%</td>
</tr>
</tbody>
</table>
What Really Causes Readmissions at Seton
Top 18 Indicators

New Insights Uncovered by Combining Content and Predictive Analytics

- LVEF and Smoking are significant indicators of CHF but not readmissions
- Assisted Living and Drug and Alcohol Abuse emerged as key predictors (only found in unstructured data)
- Many predictors are dependent on “Histories”

18. Jugular Venous Distention Indicator
17. Paid by Medicaid Indicator
16. Immunity Disorder Disease Indicator
15. Cardiac Rehab Admit Diagnosis with CHF Indicator
14. Lack of Emotion Support Indicator
13. Self COPD Moderate Limit Health History Indicator
12. With Genitourinary System & Endocrine Disorders
11. Heart Failure History
10. High BNP Indicator
9. Low Hemoglobin Indicator
8. Low Sodium Level Indicator
7. Assisted Living (NLP only)
6. High Cholesterol History
5. Presence of Blood Diseases in Diagnosis History
4. High Blood Pressure Health History
3. Self Alcohol / Drug Use Indicator (Cerner + NLP)
2. Heart Attack History
1. Heart Disease History
The Impact of Readmissions at Seton

CHF Patient X – What Happened?

The following slides will walk through a true case from Seton’s data about the predictive model alerting the propensity of readmission as well as NLP Clinical Info offered additional dimension of it.

April 18, 2009

Individual Patient Data at Each Encounter (Patient X)

- Alcohol Abuse History:
- Drug Abuse History:
- Smoking Amount:
- N/A
- Living Arrangement:
- 98%
- N/A
- N/A

Description of Model Serial Number:

1. Jugular Venous Distention Indicator
2. Immunity Disorder Disease Indicator
3. Cardiac Rehab Admit Diagnosis with CHF Indicator
4. Lack of Emotion Support Indicator
5. Self COPD Moderate Limit Health History Indicator
6. With genitourinary system & Endocrine disorders
7. Heart Failure History
8. High BNP Indicator
9. Low Hemoglobin Indicator
10. Low Sodium Level Indicator
11. Assisted Living from ICA Extract
12. High Cholesterol History
13. Presence of diseases of the blood in diagnosis history
14. High Blood Pressure Health History
15. Self Alcohol/Drug Use Indicator (Cancer + ICA)
16. Heart Attack History
17. Heart Disease History

Patient Population Monitoring Clinical and Operational Data
The Impact of Readmissions at Seton

CHF Patient X – What Happened?

Patient X was readmitted the 5th time after 26 days with additional risk factors. It surfaced that there was lack of emotional support plus Patient X had taken up smoking again as well as alcohol abuse.

98% 24 days 98% 8 days 96% 144 days 95% 44 days 96% 26 days 100%


Individual Patient Data at Each Encounter (Patient X @ Dec 20, 2009)

Description of Model Serial Number
18. Jugal Venous Distention Indicator
16. Immunity Disorder Disease Indicator
15. Cardiac Rehab Admit Diagnosis with CHF Indicator
12. With genitourinary system & Endocrine disorders
10. High BMP Indicator
9. Low Hemoglobin Indicator
7. Assisted Living from ICA Extract
6. High Cholesterol History
5. Presence of diseases of the blood in diagnosis history
4. High Blood Pressure Health History
3. Self Alcohol/Drug Use Indicator (Cancer + ICA)
2. Heart Attack History
1. Heart Disease History
The Impact of Readmissions at Seton

CHF Patient X – What Happened?

Summary of Key Readmission Risk Factors for Patient X

- **Possible Intervention Factors:** High Cholesterol, Low Sodium, Emotional Support, High Blood Pressure
- **Other Factors:** Paid by Medicaid, History: COPD, Heart Disease and Heart Failure
- A number of the top 18 factors were not available from the data at each encounter including the top predictor (Jugular Venous Distention Indicator)
IBM Content and Predictive Analytics for Healthcare

How can you apply this?

IBM Watson for Healthcare
- Confirm hypotheses or seek alternative ideas with confidence based responses from learned knowledge*

Raw Information
- Unstructured Data
  - (EMR Clinical Documentation: History and Physical, Discharge Summary, Echocardiogram.)
- Structured Data
  - (Cost Data, Admission History, Procedure History, EMR Clinical Events)

IBM Content and Predictive Analytics
- Content Analytics
  - Natural Language Processing
  - Medical Fact and Relationship Extraction (Annotation)
  - Trend, Pattern, Anomaly, Deviation Analysis
- Predictive Analytics
  - Predictive Scoring and Probability Analysis

Analyzed and Visualized Information

Dynamic Multimode Interaction
- Search and Visually Explore (Mine)
- Monitor, Dashboard and Report (Cognos BI)
- Question and Answer*
- Custom Solutions

Health Integration Framework
- Data Warehouse and Model
- Master Data Management
- Advanced Case Management
- Business Analytics

Partners (HLI)
- Specialized Research

*www.ibm.com/jstart • solutions start here. •
Thank You

jStart Emerging Technology
www.ibm.com/jstart

Watson
www.ibm.com/watson
What can I do with this data?
Visualizing the Results: Readmissions Dashboard

Cognos dashboard reporting system can help in monitoring the key clinical, operational and financial metrics. More importantly, being able to track down the top priority cases for case management.

1. Clinical Statistics: admission count, readmission count and readmission rate
2. Operational Statistic: Counts of different length of stay periods
3. Financial Statistic: Total direct cost by total admission and by readmission
4. Mortality: mortality rate
5. Average length of stay
6. Average direct cost by total admission and by readmission only
7. PA Model Score: Distribution of propensity of readmission
What can I do with this data?

Visualizing the Results: Readmissions Dashboard

Managing the follow-up cases through Cognos
Each discharged patient is scored by the predictive model to anticipate the propensity of readmission. The pool of patient is divided into 10 groups according to model risk score. The colored bar represents the average model score (0-100) of each group. The higher the model score, the higher the priority of case management.

- The colored line above the colored bar represents the distribution of encounter count in each risk group (colored bar).
- The height of the colored bars represents the average model score of the group. The higher the model score, the higher the propensity of readmission. Case manager can start to focus on the early intervention effort from the high score groups.
What can I do with this data?

Visualizing the Results: Readmissions Dashboard

Steps to Navigate through Cognos Dashboard

1. Select the date range for targeted focus group.
2. Click on colored bar to find out the encounters/patients belong to that group.
3. Select the case to follow up.
4. The individual risk profile dashboard pops up on next page.
What can I do with this data?

Visualizing the Results: Readmissions Dashboard

Components of Individual Risk Profile Dashboard

1. **Individual Risk Profile**: It displays the risk factors of patient that are flagged by predictive model.
   - i. Red line is the reference line for high risk factors.
   - ii. Orange line is the reference line for medium risk factors.

2. **Description of model variable serial number on x-axis of chart on (1)**

3. **Pie chart of count of risk factors**: Gives quick overview of number of high/medium/low risk factors to follow up with.

4. **Encounter Profile**: Provides basic profile of the encounter for case manager to follow up.

5. **Hyperlink to patient's longitudinal record**

---

18. Jugular Venous Distention Indicator
17. Paid by Medicaid Indicator
16. Immunity Disorder Disease Indicator
15. Cardiac Rehab Admit Diagnosis with CHF Indicator
14. Lack of Emotion Support Indicator
13. Self COPD Moderate Limit Health History Indicator
12. With gastrointestinal system & Endocrine disorders
11. Heart Failure History
10. High BNP Indicator
9. Low Hemoglobin Indicator
8. Low Sodium Level Indicator
7. Assisted Living From 1CA Extract
6. High Cholesterol History
5. Presence of diseases of the blood in diagnosis history
4. High Blood Pressure Health History
3. Self Alcohol/Drug Use Indicator (Cerner + ICA)
2. Heart Attack History
1. Heart Disease History

**Encounter Profile**
- **Encounter ID**: 1111111111
- **MRN**: 98
- **Admitted Facility**: UMCA
- **Model Score**: 98
- **Marital Status**: Married
- **Gender**: M
- **Ethnicity**: White

www.ibm.com/jstart • solutions start here.
Once clicked on the hyperlink of encounter, it will lead to options of viewing detail medical cost and clinical information.

1. Screen shot of Details of medical cost Information
2. Screen shot of details of clinical information (including NLP results)
IBM Content Analytics
Medical Text Analysis Implementation

- REST Web Services
  - Real-time NLP
  - Search
  - Administration

- Document Server (UIMA Pipeline)
  - ICA Annotators
  - Medical Annotators

- Extended ICA JDBC Crawler

- ICA Crawlers

- ICA Exporter

- IBM InfoSphere Optim Data Redaction

- ICA-LanguageWare Resource Workbench

- Apache Lucene Search Engine
  - Lucene Index
  - ICA-Text Miner Web Application

- Clinical Data Warehouse
  - Pathology Reports
  - Discharge Summary Reports
  - Echocardiogram Reports

- Medical Terminology
  - Health Language Inc. Language Engine
  - Custom Health Care Terminology
  - SNOMED, RxNorm, ICD-9, ICD-10, CPT-4

- Medical Terminology
  - Medical Terminology
  - SNOMED, RxNorm, ICD-9, ICD-10, CPT-4

- Clinical Data Warehouse

- Medical Text Analysis Implementation